

# The number of non-singleton blocks in $\Lambda$ -coalescents with dust

Nic Freeman\*

November 8, 2011 (*draft*)

## Abstract

Let  $\Lambda$  be a finite measure on  $[0, 1]$  and set  $\mu^n = \int_0^1 x^n \Lambda(dx)$ . Recall that a block  $b$  in a partition of  $\mathbb{N}$  is said to be a singleton if  $|b| = 1$ . It is known that the blocks of the  $\Lambda$ -coalescent are either singletons or infinite sets which comprise a non-trivial proportion of the total population. Let  $N_t^s$  and  $N_t^a$  denote the number of singleton blocks and non-singleton blocks, respectively, in the  $\Lambda$ -coalescent at time  $t > 0$ .

Pitman (1999) proved that  $N_t^s = \infty$  if  $\mu^{-1} < \infty$ , and that  $N_t^s = 0$  if  $\mu^{-1} = \infty$ . Schweinsberg (2000) gave a necessary and sufficient condition for  $N_t^a + N_t^s$  to be finite. Hence, when  $\mu^{-1} = \infty$ , Schweinsberg's result determines if  $N_t^a$  is finite or infinite. In this paper we complete the picture and show that, when  $\mu^{-1} < \infty$ , a third dichotomy occurs; if  $\mu^{-2} = \infty$  then  $N_t^a = \infty$ , and if  $\mu^{-2} < \infty$  then  $N_t^a < \infty$ .

Our proof uses the connection between  $\Lambda$ -coalescents and stochastic flows of bridges, which was established by Bertoin and Le Gall (2003).

## 1 Introduction

The  $\Lambda$ -coalescent is a Markov process, whose state space is the set  $\mathcal{P}_{\mathbb{N}}$  of partitions of  $\mathbb{N}$ . It can be thought of a system of particles, which start out separated and coagulate together over time. The  $\Lambda$ -coalescent generalizes the well known coalescent of Kingman (1982), and was introduced by Donnelly and Kurtz (1999), Pitman (1999) and Sagitov (1999) (all of which appeared independently). The name ‘ $\Lambda$ -coalescent’ comes from the formulation of Pitman (1999), which we now describe.

Let  $\mathcal{P}_n$  denote the set of partitions of  $\{1, 2, \dots, n\}$ . Let  $\iota_n$  denote the natural restriction map  $\iota_n : \mathcal{P}_{\mathbb{N}} \rightarrow \mathcal{P}_n$ , which is defined by simply removing all elements  $m \in \mathbb{N} \setminus \{1, \dots, n\}$  from a partition of  $\mathbb{N}$ . For example,  $\iota_3(\{(1, 2, 6), (3, 5), (4), \dots\}) = \{(1, 2), (3)\}$ . Let  $1^{\mathbb{N}}$  denote the partition of  $\mathbb{N}$  into singletons.

If  $\pi$  is a partition of  $\mathbb{N}$ , then each element of  $\pi$  is known as a *block*. We write  $n \stackrel{\pi}{\sim} m$  to mean that  $n$  and  $m$  are in the same block of  $\pi$ . If  $\pi = \{b_1, b_2, \dots, b_l\}$  (resp.  $\pi = \{b_i; i \in \mathbb{N}\}$ ) and  $I \subseteq \{1, \dots, l\}$  (resp.  $I \subseteq \mathbb{N}$ ) then the partition obtained from  $\pi$  by *merging*  $\{b_i; i \in I\}$  is given by  $\{b_i; i \notin I\} \cup \{\cup_{i \in I} b_i\}$ .

**Definition 1.1** *Let  $\Lambda$  be a finite measure on  $[0, 1]$ . The  $\Lambda$ -coalescent is a  $\mathcal{P}_{\mathbb{N}}$ -valued Markov process  $(\Pi_t)_{t \geq 0}$  such that, for all  $n \in \mathbb{N}$ ,  $\Pi_t^{(n)} = \iota_n(\Pi_t)$  is a  $\mathcal{P}_n$ -valued Markov chain with initial state  $1^{\mathbb{N}}$  and the following dynamics: Whenever  $\Pi_t^{(n)}$  is a partition consisting of  $i$  blocks, the rate at which any  $k$ -tuple of blocks merges is*

$$\lambda_{i,k} = \int_0^1 x^{k-2} (1-x)^{i-k} \Lambda(dx), \quad (1.1)$$

*independently of all other  $k$ -tuples.*

---

\*Mathematical Institute, Oxford University. Email: [nicholas.freeman@st-annes.ox.ac.uk](mailto:nicholas.freeman@st-annes.ox.ac.uk)

It is natural to regard the  $\Lambda$ -coalescent as a system of particles (one for each  $n \in \mathbb{N}$ ), which start out separated and which merge together over time. The precise formulation of Definition 1.1 is due to Pitman (1999).

If  $\Lambda(\{0\}) = 0$ , the formula (1.1) is more intuitively written as  $\lambda_{b,k} = \int_0^\infty x^k(1-x)^{i-k}x^{-2}\Lambda(dx)$ . The term  $x^{-2}\Lambda(dx)$  corresponds to a measure controlling the rate at which a proportion  $x \in (0, 1]$  of the blocks currently present merge to form a new block. When this occurs we call it a *coagulation event*. The remaining ‘binomial’ term  $x^k(1-x)^{i-k}$  says that, of the first  $i$  blocks, each block chooses independently whether to become part of the new block, or remain alone (with probabilities  $x$  and  $1-x$  respectively).

**Remark 1.2** *The terminology ‘coagulation event’ comes from thinking of the  $\Lambda$ -coalescent as a system of particles. The particles, one for each  $n \in \mathbb{N}$ , start out separated and coagulate together over time; a merger of blocks corresponds to a group of particles coagulating, forming a single block/particle.*

If  $\Lambda(\{1\}) > 0$ , then corresponding events occur at rate  $\Lambda(\{1\})$ , which coagulate the whole population into a single block. From a theoretical point of view, this adds no extra complexity and serves only to obfuscate the behaviour which we are trying to capture. Results concerning  $\Lambda$ -coalescents for which  $\Lambda(\{1\}) = 0$  are easily extended to general  $\Lambda$ , by simply superimposing the extra coagulation events. With this in mind:

**Remark 1.3** *In this article, without further comment, we consider only  $\Lambda$  for which  $\Lambda(\{1\}) = 0$ .*

If  $f : \mathbb{N} \rightarrow \mathbb{N}$  is a bijection, and  $\pi \in \mathcal{P}_n$  (where  $(b_i)$  is possibly a finite sequence), then we define  $f(\pi) = \{f(b); b \in \pi\}$ , where  $f(b) = \{f(k); k \in b\}$ . Recall that a (random) partition  $\pi \in \mathcal{P}_\mathbb{N}$  is said to be *exchangeable* if, for any permutation  $\sigma$  of  $\mathbb{N}$ ,  $\sigma(\pi)$  has the same distribution as  $\pi$ . Pitman (1999) showed that, for any  $t > 0$ ,  $\Pi_t$  is an exchangeable partition.

The following lemma collects together some results, due to Kingman, which can be found in Section 2.2 of Bertoin (2006). For a finite set  $A$ ,  $|A|$  denotes the cardinality of  $A$ . If  $A$  is infinite then we set  $|A| = \infty$ .

**Lemma 1.4** *Let  $\pi$  be an exchangeable partition of  $\mathbb{N}$ . Then, almost surely, for every block  $b$  of  $\pi$ , the limit*

$$\text{freq}(b) = \lim_{k \rightarrow \infty} \frac{|b \cap \{1, \dots, k\}|}{k}$$

*exists. The quantity  $\text{freq}(b)$  is known as the asymptotic frequency of  $b$ . It holds that  $\sum_{b \in \pi} \text{freq}(b) \leq 1$ , and also that*

$$\lim_{k \rightarrow \infty} \frac{|\{n = 1, \dots, k; \{n\} \in \pi\}|}{k} \stackrel{\text{a.s.}}{=} 1 - \sum_{b \in \pi} \text{freq}(b)$$

*Further, almost surely, if  $\text{freq}(b) = 0$  then  $b$  is a singleton. Almost surely,  $\pi$  has no singletons if and only if  $\sum_{b \in \pi} \text{freq}(b) = 1$ .*

**Remark 1.5** *For a general subset  $b \subseteq \mathbb{N}$ , there is no reason for the limit  $\text{freq}(b)$  to exist.*

Let us introduce some terminology, in the notation of the above lemma. If  $\text{freq}(b) > 0$ , then the elements of  $b$  comprise a non-zero proportion  $\text{freq}(b)$  of  $\mathbb{N}$ , and we refer to  $b$  as an *atomic block* of  $\pi$ .

This terminology reflects the idea that an atomic block corresponds to an atom of the probability measure which records the proportion of the population within each block.

Each singleton  $\{n\}$  comprises only a null proportion of  $\mathbb{N}$ , in that  $\text{freq}(\{n\}) = 0$ . The set of singletons of  $\pi$  is called the *dust* of  $\pi$ , and the elements of this set comprise a proportion of  $\mathbb{N}$  which is (possibly 0) and given by

$$\mathcal{D}(\pi) = 1 - \sum_{b \in \pi} \text{freq}(b).$$

Pitman (1999) proved the following result, which establishes a dichotomy in the behaviour of the dust. Let

$$\mu^n = \int_0^1 x^n \Lambda(dx). \quad (1.2)$$

**Theorem 1.6 (Pitman 1999)** *If  $\mu^{-1} < \infty$ , then  $\mathbb{P}[\forall t > 0, \mathcal{D}(\Pi_t) > 0] = 1$ , whereas if  $\mu^{-1} = \infty$  then  $\mathbb{P}[\forall t > 0, \mathcal{D}(\Pi_t) = 0] = 1$ .*

Let

$$N_t^s = |\{b \in \Pi_t; \text{freq}(b) = 0\}|,$$

which, in words, is the number of singletons of  $\Pi_t$ . By Lemma 1.4 (and the fact that  $N_u^s \leq N_v^s$  when  $u \geq v$ ), Theorem 1.6 implies (almost surely) that  $N_t^s = \infty$  or  $N_t^s = 0$ , and that  $\mathcal{D}(\Pi_t) > 0$  if and only if  $N_t^s = \infty$ .

A second dichotomy, this time in the behaviour of the total number of blocks of  $\Pi_t$ , was proven by Schweinsberg (2000). Let

$$\mu^* = \sum_{i=2}^{\infty} \left( \sum_{k=2}^i (k-1) \binom{i}{k} \lambda_{i,k} \right)^{-1}.$$

**Theorem 1.7 (Schweinsberg 2000)** *If  $\mu^* < \infty$  then  $\mathbb{P}[|\Pi_t| < \infty] = 1$ , whereas if  $\mu^* = \infty$  then  $\mathbb{P}[|\Pi_t| = \infty] = 1$ .*

It can be shown that  $\mu^{-1} < \infty$  implies that  $\mu^* = \infty$ ; in other words,  $\Lambda$ -coalescents with a non-empty dust component have infinitely many blocks for all time (of course, we already knew this since non-empty dust implies  $N_t^s = \infty$ ). Recall that the  $\Lambda$ -coalescent is said to *come down from infinity* if  $\mathbb{P}[\forall t > 0, |\Pi_t| < \infty] = 1$ . Theorem 1.7 shows that the  $\Lambda$ -coalescent comes down from infinity if and only if  $\mu^* < \infty$ .

The main result of this article is to establish a third (and, in some sense, final) dichotomy which occurs for  $\Lambda$ -coalescents. Let

$$N_t^a = |\{b \in \Pi_t; \text{freq}(b) > 0\}|,$$

which is the number of atomic blocks of  $\Pi_t$ . Note that

$$N_t^a + N_t^s = |\Pi_t|.$$

If  $\mu^{-1} = \infty$  then (see above),  $\Pi_t$  has no singletons, and Theorem 1.7 tells us that  $N_t^a = \infty$  when  $\mu^* = \infty$  and  $N_t^a < \infty$  when  $\mu^* < \infty$ . The behaviour of  $N_t^a$ , in the case  $\mu^{-1} < \infty$ , is described by the following theorem.

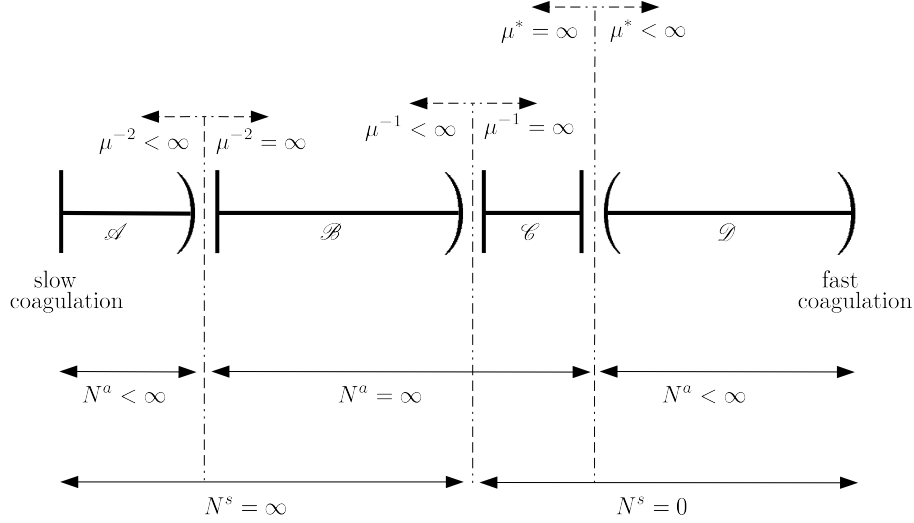


Figure 1: The qualitative changes in the behaviour of the  $\Lambda$ -coalescent are shown. The behaviour of the  $\Lambda$ -coalescent is determined solely by the behaviour of the measure  $\Lambda$  at  $0+$ , that is, by the rate of small coagulation events. Moving from left to right, across the picture, the rate of these coagulation events increases. The resulting four behaviours are labelled  $\mathcal{A}$  to  $\mathcal{D}$ .

**Theorem 1.8** Suppose that  $\mu^{-1} < \infty$ . If  $\mu^{-2} = \infty$  then  $\mathbb{P}[\forall t > 0, N_t^a = \infty] = 1$ , whereas if  $\mu^{-2} < \infty$  then  $\mathbb{P}[\forall t > 0, N_t^a < \infty] = 1$ .

With Theorems 1.6-1.8 in hand, the qualitative behaviour of  $N^a$  and  $N^s$  can be completely classified, as in Figure 1.

It is surprising that Theorem 1.8 was not discovered until now; in fact we were unable to find even a single example in the literature of a  $\Lambda$ -coalescent which was shown to have both  $\mathcal{D}(\Pi_t) > 0$  and  $N_t^a = \infty$ . The question ‘are there  $\Lambda$ -coalescents with  $\mathcal{D}_t > 0$  and  $N_t^a = \infty$ ’ naturally emerged out of the analysis in Freeman (2011), see Remark 1.9. We will discuss our proof of Theorem 1.8 in Section 1.1. See Example 1.10 for an application of Theorem 1.8 to  $\beta$ -coalescents.

**Remark 1.9** In Freeman (2011), a natural extension of the  $\Lambda$ -coalescent is defined in a spatial continuum, and its behaviour is classified in a similar manner to that of Theorems 1.6-1.8. It is shown there that the introduction of space enriches the behaviour of the  $\Lambda$ -coalescent. For example, if  $p_t$  is the probability that the coalescent of Freeman (2011) has finitely many blocks at time  $t$ , it is sometimes the case that, for some deterministic  $t_0 \in (0, \infty)$ ,  $p_t > 0$  for  $t < t_0$  and  $p_t = 0$  for  $t \geq t_0$ .

**Example 1.10 ( $\beta$ -coalescents)** The  $\Lambda$ -coalescent where  $\Lambda(dx)$  has the  $\beta(2 - \alpha, \alpha)$  distribution, for  $\alpha \in (0, 2)$ , is known as the  $\beta$ -coalescent with parameter  $\alpha$ . The  $\beta$ -coalescents are one of the most tractable families of  $\Lambda$ -coalescents, see for example Berestycki et al. (2008).

The first example of a (non-Kingman)  $\Lambda$ -coalescent, which was introduced in Bolthausen and Sznitman (1998), corresponds to case of a  $\beta$ -coalescent with  $\alpha = 1$ . In this case  $\Lambda$  is the uniform probability measure on  $[0, 1]$ . The case  $\alpha = 2$  is, by convention (or as the natural limiting case) Kingman’s coalescent, where  $\Lambda$  is a point mass at 0.

Label the different behaviours of the  $\Lambda$ -coalescent as  $\mathcal{A}$  to  $\mathcal{D}$ , as in Figure 1. Some easy estimates show that

- For  $\alpha \in (0, 1)$ , behaviour  $\mathcal{B}$  occurs
- For  $\alpha = 1$ , behaviour  $\mathcal{C}$  occurs.
- For  $\alpha \in (1, 2]$  behaviour  $\mathcal{D}$  occurs.

Behaviour  $\mathcal{A}$  does not occur amongst  $\beta$ -coalescents. However, it is easy to construct  $\Lambda$ -coalescents for which behaviour  $\mathcal{A}$  does occur; for example  $\Lambda(dx) = x^2 dx$ .

## 1.1 Outline of the proof of Theorem 1.8

The case  $\mu^{-2} < \infty$  is well known, and corresponds to the case where coagulation events (of  $\Pi_t$ ) occur only at finite rate. Each coagulation event produces at most one new atomic block, and hence the total number of atomic blocks is always finite. We give an argument based on this intuition in Section 3.1.

Let us now explain why the other case,  $\mu^{-2} = \infty$ , is not so obvious. In the case  $\mu^{-2} = \infty$ , it is immediate that (in some limiting sense) infinitely many coagulation events occur during all non-trivial time intervals. Since  $\mu^{-1} < \infty$  (as an assumption of the theorem), at all times  $\Pi_t$  contains a dust component with  $\mathcal{D}(\Pi_t) > 0$ . Hence, each coagulation event takes a proportion of singletons out of the dust component, and merges them into a single new block. The catch is that each coagulation event also takes a proportion of the atomic blocks, and merges them too into the new block. Therefore, the argument from the case  $\mu^{-2} < \infty$  does not work in reverse; it is not obvious that having infinitely many coagulation events is enough to guarantee infinitely many atomic blocks.

Note that, although the singletons decrease in number at each coagulation event, the atomic blocks can both increase and decrease in number. This is easily seen by considering  $\Pi_t^n$ . Therefore, it could potentially be the case that  $N_t^a < \infty$  and  $N_{t'}^a = \infty$  occurred at two different (random or deterministic) times  $t, t' \in (0, \infty)$ . With this observation in mind, it is not obvious that a result as clear cut as Theorem 1.8 holds.

But, it is natural to suspect that for at least some choices of  $\Lambda$ , for at least some  $t > 0$ , we might have  $\mathcal{D}_t > 0$  and  $N_t^a = \infty$ . Let us think, therefore, about how we might go about identifying such cases.

One option is to consider the restrictions  $\Pi_t^n$ , and hope to show that as  $n \rightarrow \infty$ , the number of non-singleton blocks of  $\Pi_t^n$  tends to infinity as  $n \rightarrow \infty$ . There is some difficulty involved in getting direct estimates on the behaviour of  $\Pi_t^n$ , especially in the general case where the behaviour of  $\Lambda$  might vary wildly near 0. We choose not to attempt these calculations, noting that there is already a wealth of literature offering more sophisticated ways of analysing the  $\Lambda$ -coalescent.

**Remark 1.11** *For example, Pitman (1999), Bertoin and Le Gall (2003), and Birkner et al. (2005) offer three genuinely different ways in which to represent the  $\Lambda$ -coalescent. A further representation (for an important special case) is given in Berestycki et al. (2008). We refer the reader to Berestycki (2009) for further references.*

In general, the  $\Lambda$ -coalescent is an infinite rate object and, in the spirit of the above paragraph, it is sensible to use more tractable finite rate approximations. As we saw in Definition 1.1, one way to achieve this is by looking at only a finite subset of the particles (i.e. blocks) which are active in the system. A second option is to reduce the total rate of the system to something finite, and

consider approximation by a sequence of coalescents, each acting on  $\mathbb{N}$ , but with only finite rate coagulation events.

To be precise, we could consider a sequence of  $\Lambda$ -coalescents, with

$$\int_0^1 x^{-2} \Lambda^{(n)}(dx) < \infty,$$

such that  $\Lambda^{(n)} \rightarrow \Lambda$ , in some sense. This is borne out in the work of Bertoin and Le Gall (2003), and the corresponding construction of the  $\Lambda$ -coalescent will be recapped in Section 2. Using this construction, we have excellent global control over the number of atomic blocks, and we will be able to get natural estimates on the number and corresponding asymptotic frequencies of the atomic blocks which appear in the finite rate approximations.

However, there is still one major problem to solve; the construction from Bertoin and Le Gall (2003) produces the  $\Lambda$ -coalescent as a limiting object, in the sense of finite dimensional distributions. Consequently, and in contrast to Definition 1.1, the theory of Bertoin and Le Gall (2003) allows us to make estimates which are uniform in space (i.e. valid for arbitrarily many particles), but at the expense of uniformity in time.

A minor technicality arises in the limit taking, since a partition with no atomic blocks, is close (in a suitable metric space sense) to a partition with a large (possibly infinite) number of very small atomic blocks. Therefore, some degree of care is needed to make sure the blocks which appear in our finite rate approximations really do appear in the limit. As one might expect, the issue is overcome by (loosely speaking) taking an arbitrary  $\epsilon > 0$  and working with blocks  $b \in \Pi_t$  that have  $\text{freq}(b) \geq \epsilon > 0$ .

Since the limit is taken in terms of finite dimensional distributions, we arrive at only the result  $N_t^a = \infty$  for fixed, deterministic, times  $t \in (0, \infty)$ . It could, potentially, still be the case that  $N_T^a < \infty$  at some random time  $T$  or, even worse, on a random null (but potentially dense) subset of  $(0, \infty)$ .

We regain global control in time through the following observation. Fix deterministic  $t' > 0$  and consider the set  $S$  of atomic blocks of  $\Pi_{t'}$ . Regard each of these blocks as a singletons in the initial state of a new coalescent, and define the evolution of our new coalescent via the coagulation induced from  $(\Pi_{t'})_{t > t'}$ . Let us name this new coalescent, which acts on  $S$ , as  $\tilde{\Lambda}_S$ . It is not hard to see that  $\tilde{\Lambda}_S$  is a  $\Lambda$ -coalescent, with the same coagulation rates as  $\Pi$ . Since  $\mu^{-1} < \infty$ ,  $\Pi$  does not come down from infinity, and hence neither does  $\tilde{\Lambda}_S$ . Hence,  $N_t^a = \infty$  for all  $t \geq t'$ . Taking  $t'$  to be arbitrarily small completes the proof of Theorem 1.8.

## 2 Flows of bridges

In this section we recap some well known results of Bertoin and Le Gall (2003), namely the construction of  $\Lambda$ -coalescents using flows of bridges. From now on we switch to using the measure

$$\nu(dx) = \frac{1}{x^2} \Lambda(dx),$$

since it will be intuitively clearer to see the  $\Lambda$ -coalescent in terms of  $\nu$ , rather than  $\Lambda$ .

**Remark 2.1** *In terms of  $\nu$ , the conditions for Theorem 1.8 to apply are that  $\nu(\{0, 1\}) = 0$ , and  $\int_0^1 x \nu(dx) < \infty$ . Note that the apparently extra assumption that  $\nu(\{0\}) = 0$  is implied by*

$\int_0^1 x^{-1} \Lambda(dx) < \infty$ , via the convention that  $1/0 = \infty$ . The dichotomy stated in Theorem 1.8 then rests on whether  $\nu([0, 1])$  is finite or infinite.

The condition for the existence of a  $\Lambda$ -coalescent corresponding to  $\nu$  is that  $\int_0^1 x^2 \nu(dx) < \infty$ .

In keeping with our new terminology, we refer to the  $\Lambda$ -coalescent  $(\Pi_t)$  corresponding to  $\nu$  as the  $\nu$ -coalescent, but we continue to use the term  $\Lambda$ -coalescent as a name for the general family of processes.

For the remainder of Section 2, let  $\nu$  be a measure on  $[0, 1]$  such that  $\int_0^1 x^2 \nu(dx) < \infty$  and  $\nu(\{0, 1\}) = 0$ . Let  $(\Pi_t)_{t \geq 0}$  denote the  $\nu$ -coalescent.

**Remark 2.2** *In order to state the results of Bertoin and Le Gall (2003), we will first need to develop some theory relating to exchangeable partitions. For this reason, it is convenient for us to simply use Sections 2.1, 2.3 and 4.4 of Bertoin (2006) as our main reference for this section.*

**Definition 2.3** *Let  $\mathcal{P}_{\mathcal{M}}$  be the space of sequences  $s = (s_i)_{i=1}^\infty \subseteq [0, 1]$  which are such that*

$$s_1 \geq s_2 \geq \dots \geq 0 \quad \text{and} \quad \sum_{i=1}^\infty s_i \leq 1$$

*The set  $\mathcal{P}_{\mathcal{M}}$  is a compact metric space, endowed with the metric  $d(s, s') = \max\{|s_i - s'_i|; i \in \mathbb{N}\}$ .*

For an exchangeable partition  $\pi = \{b_1, b_2, \dots\}$  of  $\mathbb{N}$ , we write  $|\pi| = (\text{freq}(b_1), \text{freq}(b_2), \dots)$ . If  $\pi = \{b_1, \dots, b_n\}$  has only finitely many blocks, we set  $|\pi| = (\text{freq}(b_1), \dots, \text{freq}(b_n), 0, 0, \dots)$ . We write  $|\pi|^\downarrow$  for the (random) sequence obtained by rearranged the elements of  $|\pi|$  into decreasing order. The sequence  $|\pi|^\downarrow$  is known as the sequence of *asymptotic frequencies* of  $\pi$ . For any exchangeable partition  $\pi \in \mathcal{P}_{\mathbb{N}}$ ,  $|\pi|^\downarrow$  is an element of  $\mathcal{P}_{\mathcal{M}}$ .

The following lemma will be of use to us, later on.

**Lemma 2.4 (Proposition 2.1, Bertoin 2006)** *Let  $(s^n)$  be a sequence in  $\mathcal{P}_{\mathcal{M}}$ , where  $s^n = (s_i^n)_{i=1}^\infty$  and  $s \in \mathcal{P}_{\mathcal{M}}$ . Then  $d(s^n, s) \rightarrow 0$  if and only if for all  $i \in \mathbb{N}$ ,  $s_i^n \rightarrow s_i$ .*

Let  $L^2[0, 1]$  denote the space of (equivalence classes of) square integrable functions  $f : [0, 1] \rightarrow [0, 1]$ , equipped with the usual  $L^2$  metric,

$$d_{L^2}(f, g) = \left( \int_0^1 |f(x) - g(x)|^2 dx \right)^{1/2}.$$

**Definition 2.5** *Let  $s = (s_i)_{i=1}^\infty \in \mathcal{P}_{\mathcal{M}}$ , and let  $s_0 = 1 - \sum_{i=1}^\infty s_i$ . Let  $(U_i)$  be an infinite sequence of independent uniform random variables on  $[0, 1]$ . Any process which has the same distribution (in  $L^2[0, 1]$ ) as*

$$b_s(y) = s_0 y + \sum_{i=1}^\infty s_i \mathbb{1}\{U_i \leq y\}$$

*is known as an  $s$ -bridge (or, when  $s$  is not specified, a bridge). By definition, when we use more than one bridge at once, we use a different independent sequence of uniform random variables for each bridge.*

*If  $s$  is random, then by definition we choose  $(U_i)_1^\infty$  to be independent of  $s$ , and we also say  $b_s$  is an  $s$ -bridge.*

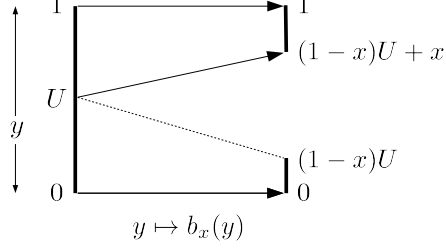


Figure 2: The simple bridge  $b_x(y) = (1-x)y + x\mathbb{1}\{U \leq y\}$ .

**Definition 2.6** If  $s = (x, 0, 0, \dots)$  then we say the bridge  $b_s$  is a simple bridge, and write

$$b_s(y) = b_x(y) = (1-x)y + x\mathbb{1}\{U \leq y\}.$$

Here  $U$  is a uniform random variable on  $[0, 1]$ , and  $U$  is independent of  $s$ .

**Definition 2.7** A bridge  $b_s$  where  $s = (s_1, \dots, s_n, 0, 0, \dots)$  is said to be a finite bridge. In the case where  $s$  is random,  $n$  may be random and dependent on  $s$  (but not on the  $(U_i)$ ).

Note that a bridge is a right continuous, strictly increasing function. It is easily checked that the composition of finitely many bridges is a bridge, and that the composition of finitely many finite bridges is a finite bridge. But note that the composition of finitely many simple bridges is only a finite bridge. It will be convenient to picture our bridges as stochastic flows, rather than functions, as indicated in Figure 2.

For an increasing function  $f : [0, 1] \rightarrow [0, 1]$ , let  $f^{-1}$  denote the right continuous inverse of  $f$ , that is

$$f^{-1}(y) = \inf\{z \in [0, 1]; f(z) > y\}.$$

The following two Lemmas summarise the connection between exchangeable partitions and bridges, which is developed in Section 4.4 of Bertoin (2006).

**Lemma 2.8** Let  $b_s$  be a bridge, and let  $(V_i)_{i=1}^\infty$  be a sequence of uniform random variables on  $[0, 1]$ , independent of each other and of  $s$ . Define a partition  $\pi$  of  $\mathbb{N}$  by

$$i \stackrel{\pi}{\sim} j \iff b_s^{-1}(V_i) = b_s^{-1}(V_j).$$

Then  $\pi$  is an exchangeable random partition. Further,  $\mathcal{D}(\pi) = s_0 = 1 - \sum_{i=1}^\infty s_i$ , and  $|\pi|^\downarrow = (s_i)_{i=1}^\infty$ .

**Lemma 2.9** Every exchangeable partition  $\pi$  of  $\mathbb{N}$  is equal in law to the exchangeable partition constructed by Lemma 2.8 from  $s = |\pi|^\downarrow$ .

We extend the definition of  $\mathcal{D}$  in the natural way, as follows. If  $b$  is a bridge,  $\mathcal{D}(b_s)$  denotes the size of the dust in the corresponding exchangeable partition.

Bertoin and Le Gall (2003) showed that, with a careful choice of bridges, it is possible to use Lemma 2.8 to construct the  $\Lambda$ -coalescent (as a limit, in terms of finite dimensional distributions). To this end, let  $M$  be a Poisson point process with points  $(t, x) \in (0, \infty) \times [0, 1]$  and intensity measure

$$dt \otimes \nu(dx).$$



Here  $dt$  denotes Lebesgue measure. We now make several ‘without loss of generality’ assumptions, since it will be convenient to have several almost sure properties of  $M$  as ‘sure’ properties of  $M$ .

Since  $\nu(\{0, 1\}) = 0$ , almost surely  $M$  is contained in  $(0, \infty) \times (0, 1)$ , so without loss of generality we will assume that  $M \subseteq (0, \infty) \times (0, 1)$ . It follows from  $\sigma$ -finiteness of  $dt \otimes \nu(dx)$  that, almost surely, for all  $t \in (0, \infty)$ , at most one point of  $M$  has  $t$  as its first coordinate. So, without loss of generality, we will assume this is the case for all realizations of  $M$ . Similarly, almost surely, for each  $x \in (0, 1)$ , at most one point of  $M$  has  $x$  as its second coordinate; without loss of generality we assume this, too, is the case for all realizations of  $M$ . For all  $n \in \mathbb{N}$ , since  $\nu((1/n, 1])$  is a finite measure,  $M \cap (0, \infty) \times (1/n, 1)$  is almost surely a finite set. So, finally, without loss of generality we assume that both  $M \cap (0, \infty) \times (1/n, 1)$  is finite, for all realizations of  $M$  and all  $n, t, t'$ .

If  $\nu([0, 1]) = \infty$ , it follows that for all  $0 \leq t < t'$ ,  $M \cap (t, t'] \times (0, 1)$  is almost surely a countably infinite set. If this is the case, without loss of generality, we assume also that  $M \cap (t, t'] \times (0, 1)$  is countably infinite for all  $0 \leq t < t'$ . If  $\nu([0, 1]) < \infty$  then for all  $0 \leq t < t'$ ,  $M \cap (t, t'] \times (0, 1)$  is almost surely a finite set, and, without loss of generality, if  $\nu([0, 1]) < \infty$  we assume this is so. We adopt the convention that a Poisson random variable  $X \sim \text{Poisson}(\infty)$  simply means  $X = \infty$ .

Fix  $0 \leq t < t' < \infty$ . Set  $N_\infty = |M|$ , and note that by the above,

$$N_\infty \sim \text{Poisson}((t' - t)\nu([0, 1])).$$

Enumerate  $M \cap (t, t'] \times (0, 1)^2$  such that

$$M \cap (t, t'] \times (0, 1) = \{(t_i, x_i); i = 1, \dots, N_\infty\}$$

where, for all  $i$ ,  $x_i < x_{i+1}$ . For each  $m \in \mathbb{N}$  and  $0 \leq t < t'$ , define

$$N_m = \#(M \cap (0, 1) \times (1/m, 1))$$

and note that then  $\{(t_i, x_i); i = 1, \dots, N_m\} = M \cap (0, 1) \times (1/m, 1)$ , so as

$$N_m \sim \text{Poisson}((t' - t)\nu((1/n, 1])). \quad (2.1)$$

It is easy to see that  $N_m \leq N_{m+1}$ , and that  $N_m \rightarrow N_\infty$  as  $m \rightarrow \infty$ .

For each  $n \in \mathbb{N} \cap \{1, \dots, N_\infty\}$ , define the bijection  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  by the properties

$$\begin{aligned} \{(t_i, x_i); i = 1, \dots, n\} &= \{(t_{\sigma(i)}, x_{\sigma(i)}); i \in \mathbb{N}\} \\ \text{for all } i, \quad t_{\sigma(i)} &< t_{\sigma(i+1)} \end{aligned}$$

To avoid excessive subscripts, we do not normally write the dependence of  $\sigma = \sigma_n$  on  $n$ . Set

$$B_{t,t'}^{(n)} = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(n)}}. \quad (2.2)$$

If  $N_\infty < \infty$  and  $n \in \mathbb{N}$ ,  $n > N_\infty$ , then we define  $B_{t,t'}^{(n)} = B_{t,t'}^{(N_\infty)}$ . Note that we have defined  $B_{t,t'}^{(n)}$  for all  $n \in \mathbb{N}$ , but not for  $n = \infty$ . The limiting case  $n = \infty$  will be defined as part of Theorem 2.12. For all  $n \in \mathbb{N}$ ,  $B_{t,t'}^{(n)}$  is a finite bridge.

**Remark 2.10** *Our notation for the order of function composition is that  $(f \circ g)(x) = g(f(x))$ . This is the same as in Bertoin (2006).*

**Remark 2.11** To summarise the above, we impose an order on  $M = \{(t_i, x_i); i \in \mathbb{N}\}$  by ranking the second coordinates of elements of  $M$  in decreasing order. Then,  $\{(t_{\sigma(i)}, x_{\sigma(i)}); i \in \mathbb{N}\}$  is the first  $n$  elements of  $M$ , reordered by time coordinate.

The random function  $B_{t,t'}^{(n)}$  is the composition of the simple bridges corresponding to the first  $n$  points of  $M$ , composed according to the order of their time coordinates. Similarly, the random function  $B_{t,t'}^{(N_m)}$  is the composition, after appropriate reordering, of the bridges corresponding to  $M \cap (t, t') \times (1/n, 1)$ .

We denote the right continuous inverse of  $B_{t,t'}^{(n)}$  by  $B_{t,t'}^{(n),-1}$ .

**Theorem 2.12** *The following holds.*

1. For each  $0 \leq t < t'$ , the sequence  $B_{t',t}^{(N_m)}$  converges in law.

Denote the limit by  $B_{t,t'}^{(\infty)}$ , and its right continuous inverse by  $B_{t,t'}^{(\infty),-1}$ .

Now fix  $t > 0$ . Let  $(V_i)_{i=1}^\infty$  be a sequence of uniform random variables on  $[0, 1]$ , which are independent of  $\{B_{0,t}^{(n)}; n \in \mathbb{N} \cup \{\infty\}\}$ , and of each other. For each  $n \in \mathbb{N} \cup \{\infty\}$  and  $t > 0$  define a random partition  $\pi_t^{(n)}$  of  $\mathbb{N}$  by

$$i \stackrel{\pi_t^{(n)}}{\sim} j \iff B_{0,t}^{(n),-1}(V_i) = B_{0,t}^{-1}(V_j).$$

2. Then,  $\pi_t^{(\infty)}$  and  $\Pi_t^{(\infty)}$  have the same distribution.
3. As  $m \rightarrow \infty$ ,  $\pi_t^{(N_m)}$  converges to  $\pi_t^{(\infty)}$  in distribution.
4. As  $m \rightarrow \infty$ ,  $|\pi_t^{(N_m)}|^\downarrow$  converges to  $|\pi_t^{(\infty)}|^\downarrow$  in distribution.

PROOF: Essentially, this result is Theorem 4.3 of Bertoin (2006). Note, however, that Bertoin defines the symbol  $B^{(n)}$  slightly differently to us (see (2.2) for our definition). Rather than considering each point of  $M$  as a separate entity, Bertoin's  $B_{t,t'}^{(n)}$  is a composition of the bridges corresponding to all points of the finite set  $M \cap (t, t'] \cap (1/n, 1)$ .

The connection is that, for  $m \in \mathbb{N}$ , Bertoin's  $B_{t,t'}^{(m)}$  is our  $B_{t,t'}^{(N_m)}$ . Having realized this, statements 1 and 2 are simply restatements of facts contained within Theorem 4.3 of Bertoin (2006). Statement 3 is proved during the course of the proof of that theorem, and statement 4 follows immediately from 3, by applying Proposition 2.9 of Bertoin (2006). ■

**Remark 2.13** *Theorem 4.3 of Bertoin (2006) also states that  $B_{t,t'}$  is a stochastic flow (in some sense), but Theorem 2.12 contains all the information needed for our purposes.*

### 3 Proof of Theorem 1.8

For the remainder of the article, suppose that  $\int_0^1 x\nu(dx) < \infty$  and that  $\nu(\{0, 1\}) = 0$ . We consider the case  $\nu([0, 1]) < \infty$  in Section 3.1, and the remaining sections are devoted to the case  $\nu([0, 1]) = \infty$ .

A central concept of the proof is the idea of a hole in a finite bridge, defined as follows.

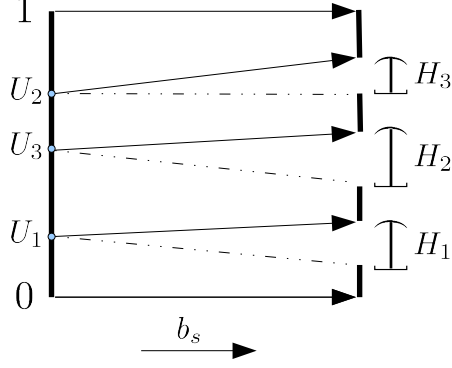


Figure 3: The holes  $H_1, H_2, H_3$  of the bridge  $b_s(y) = (1 - s_0)y + \sum_{i=1}^3 s_i \mathbb{1}\{U_i \leq y\}$ .

**Definition 3.1** A set  $H \subseteq [0, 1]$  is said to be a hole of the finite bridge  $b_s$  if  $H$  is a maximally connected component of  $[0, 1] \setminus b_s([0, 1])$ .

Since  $b_s$  is right continuous and has only finitely many discontinuities,  $b_s$  has finitely many holes, each of which is a half open interval  $H = [h_1, h_2)$ . The size of the hole  $H = [h_1, h_2)$  is given by

$$\mathcal{S}(H) = h_2 - h_1.$$

See Figure 3. Note that  $\mathcal{S}(H) > 0$  for all holes  $H$  of  $b_s$ . We say  $h_2$  is the *boundary point* of the hole  $H = [h_1, h_2)$  and write  $h_2 = \partial H$ .

Holes are a natural way of thinking about the connection between bridges and exchangeable partitions, as the following Lemma shows.

**Lemma 3.2** Let  $b_s$  be a bridge and let  $\pi$  be the exchangeable partition of  $\mathbb{N}$  constructed by Lemma 2.8 from  $b_s$ . Then there is a bijective correspondence between the holes of  $b_s$  and the atomic blocks of  $\pi$ , such that

$$\mathcal{S}(H) = \text{freq}(b)$$

(where  $b$  is an atomic block of  $\pi$  and  $H$  is a hole of  $b_s$ ).

PROOF: Recall the construction of Lemma 2.8, which is usually known as Kingman's ‘paintbox’ construction (described in Section 4.4 of Bertoin 2006). Note that each hole  $H$  of size  $x$  corresponds to a level set of  $b_s^{-1}$ , namely

$$\{y \in [0, 1]; b_s^{-1}(y) = b_s^{-1}(\partial H)\}.$$

The probability that each  $V_i$  falls into this level set is precisely  $x$ , which implies that the corresponding block in  $\pi$  has asymptotic frequency  $x$ . ■

The real advantages of thinking in terms of holes is that they provide a way of tracking the atomic blocks through the composition

$$B_{0,t}^{(n)} = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(n)}}.$$

We will show precisely how this is achieved in the coming sections.

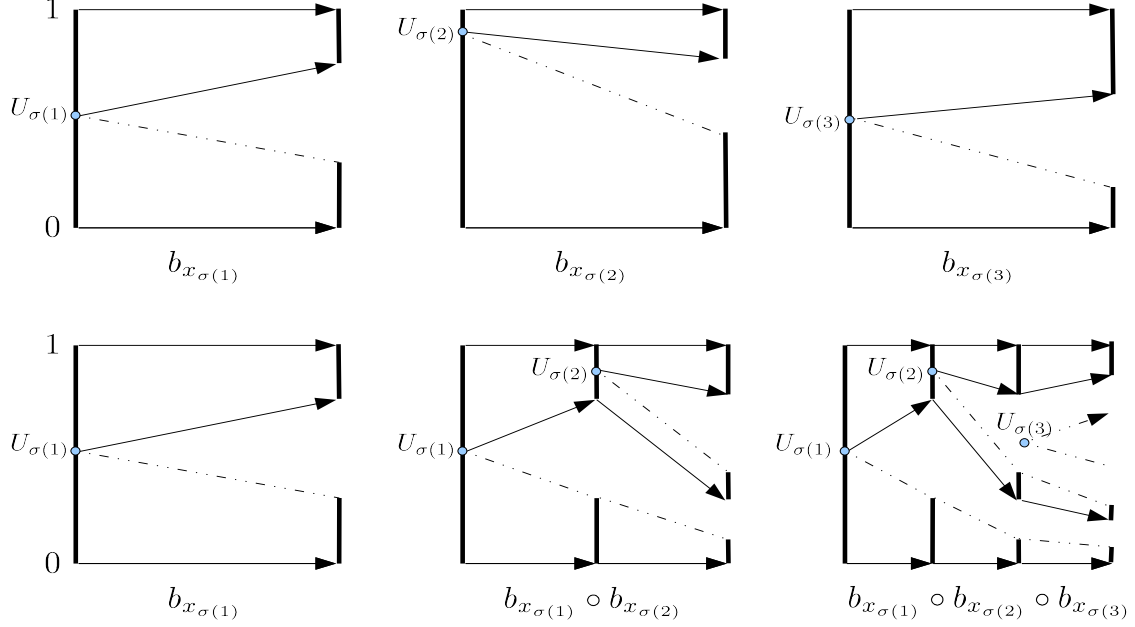


Figure 4: The composition of three simple bridges. Note that composition with  $b_{x_{\sigma(2)}}$  adds a new hole, but composition with  $b_{x_{\sigma(3)}}$  does not. Recall that  $(f \circ g)(x) = g(f(x))$ .

### 3.1 The case $\nu([0, 1]) < \infty$

The case  $\nu([0, 1]) < \infty$  is straightforward and is already well understood, see Example 19 of Pitman (1999). We will give a heuristic proof using flows of bridges (which is easily made rigorous), as a way of familiarising the reader with some of the methods which we use in Section 3.2.

Since  $\nu([0, 1]) < \infty$ ,  $N_\infty \sim \text{Poisson}(t\nu([0, 1]))$ , and in particular  $\mathbb{P}[N_\infty < \infty] = 1$ . Hence, there exists a (random)  $M \in \mathbb{N}$  such that, for all  $m \geq M$ ,

$$B_{0,t}^{(N_m)} = B_{0,t}^{(N_M)}.$$

By Theorem 2.12,  $\Pi_t$  has the same distribution as the partition associated to  $B_{0,t}^{(N_M)}$ . Now, since  $N_M \leq N_\infty < \infty$ ,  $B_{0,t}^{(N_M)}$  is a finite bridge. Setting  $\sigma = \sigma_{N_M}$ , we have

$$B_{0,t}^{(N_M)} = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(N_M)}}.$$

An illustration of the composition of finite bridges is given as Figure 4.

From Figure 4, we see that composing a new bridge (on the right) with a finite bridge will either (1) keep the number of holes constant, or (2) increase the number of holes by 1. Hence,  $B_{0,t}^{(N_M)}$  has at most  $N_M < \infty$  holes. By Lemma 3.2,  $\pi_t^{(N_m)}$  has only finite many atomic blocks, and (heuristically at least; there are technicalities to take care of at this point) the limit has the same atomic blocks as  $\pi_t^{(N_m)}$ , which finishes the argument.

### 3.2 At a single deterministic time

We now focus on the case  $\nu([0, 1]) = \infty$ , which is the new part of the content of Theorem 1.8. Recall that  $N_\infty = |M| = \infty$ , in this case. In this section we prove the required result for a single deterministic time  $t \in (0, \infty)$ . That is, we prove

**Theorem 3.3** *Let  $t > 0$ . Then  $\mathbb{P}[N_t^a = \infty] = 1$ .*

For the remainder of this section we fix some  $t > 0$ . Before we embark on the proof, we need to collect together some lemmas. For  $f : [0, 1] \rightarrow [0, 1]$ , let

$$\text{lip}^\downarrow(f) = \sup\{\kappa \in [0, \infty) ; \forall y, z \in [0, 1], |f(y) - f(z)| \geq \kappa|y - z|\}$$

be the lower Lipschitz constant of  $f$ .

**Lemma 3.4** *Let  $b_s$  be a finite bridge, where  $s = (s_i)$ . Then  $\mathcal{D}(b_s) = \text{lip}^\downarrow(b_s)$ .*

PROOF: By definition, we have  $b_s(y) = (1 - s_0)y + \sum_1^n s_i \mathbb{1}\{U_j \leq y\}$ , where  $(U_j)$  are uniform random variables on  $[0, 1]$ , independent of both each other and of  $s$ . Hence  $\text{lip}^\downarrow(b_s) = s_0$ . The result follows by Lemma 2.8.  $\blacksquare$

**Lemma 3.5** *If  $b_s$  is a finite bridge and  $b_x$  is a simple bridge then  $\mathcal{D}(b_x) = 1 - x$  and*

$$\mathcal{D}(b_s \circ b_x) = \mathcal{D}(b_s)\mathcal{D}(b_x).$$

PROOF: Again, by definition, we have  $b_s(y) = (1 - s_0)y + \sum_1^n s_i \mathbb{1}\{U_j \leq y\}$ , where  $(U_j)$  are uniform random variables on  $[0, 1]$ , independent of both each other and of  $s$ . If we write the  $U_i$  in increasing order as  $U_{\tau(i)}$ , then for each  $y, z \in [U_{\tau(i)}, U_{\tau(i+1)}]$ ,

$$|b_s(y) - b_s(z)| = s_0|y - z|. \tag{3.1}$$

The result follows immediately from (3.1) and Lemma 3.4, since composing (local) isometries corresponds to multiplying  $\text{lip}^\downarrow$ .  $\blacksquare$

We now work towards Theorem 3.3.

**Lemma 3.6** *The following hold.*

1. *For all  $n \in \mathbb{N}$ ,  $\mathcal{D}(B_{0,t}^{(n)}) \geq \mathcal{D}(B_{0,t}^{(n+1)})$ .*
2. *The limit  $D_t = \lim_{m \rightarrow \infty} \mathcal{D}(B_{0,t}^{(N_m)})$  exists almost surely.*
3.  *$D_t$  and  $\mathcal{D}(\Pi_t)$  have the same distribution.*

PROOF: Using Lemma 3.5 iteratively on (2.2),

$$\mathcal{D}(B_{0,t}^{(n)}) = \prod_{i=1}^n (1 - x_{\sigma(i)}) \tag{3.2}$$

where  $\sigma = \sigma_n$ . Statement 1 follows immediately. Hence, the limit  $D_t = \lim_{m \rightarrow \infty} \mathcal{D}(B_{0,t}^{(N_m)})$  exists almost surely (since monotone decreasing sequences converge), which proves 2. Note that in fact  $D_t = \lim_{n \rightarrow \infty} \mathcal{D}(B_{0,t}^{(n)})$  (although we will not have need of this fact).

Note that 3 concerns only the distribution  $\Pi_t$  and  $\pi_t^{(\infty)}$ . Let us write

$$B_{0,t}^{(N_m)} = b_{(s_i^m)_{i=1}^\infty},$$

so that  $|\pi_t^{(N_m)}|^\downarrow = (s_i^m)$ , and similarly write  $|\Pi_t|^\downarrow = (s_i^\infty)$ . By Theorem 2.12,  $|\pi_t^{(N_m)}|^\downarrow \rightarrow |\Pi_t|^\downarrow$  in distribution. By the compactness of  $\mathcal{P}_{\mathcal{M}}$ , the Skorohod Representation Theorem implies that we may assume (in as far as proving 3 is concerned) that  $|\pi_t^{(N_m)}|^\downarrow \rightarrow |\Pi_t|^\downarrow$  almost surely. Hence, by Lemma 2.4,  $s_i^m \rightarrow s_i^\infty$ , for all  $i \in \mathbb{N}$ . By Lemma 2.8,

$$\mathcal{D}(\Pi_t) = 1 - \sum_1^\infty s_i^\infty,$$

and by Lemma 3.4,

$$\mathcal{D}(\pi_t^{(N_m)}) = 1 - \sum_1^\infty s_i^m.$$

By dominated convergence, almost surely,  $\lim_m 1 - \sum_1^\infty s_i^m = 1 - \sum_1^\infty s_i^\infty$ , which completes the proof.  $\blacksquare$

**Remark 3.7** *A change of probability space occurs in the above proof, caused by the application of Skorohod's Representation Theorem. This change preserves the distribution of each  $\pi_t^{(N_m)}$ , and the limit  $\Pi_t$ , individually, but does not preserve their joint distribution. We will use the joint distribution in the sequel, so, outside of the above proof, we do not change our background probability space. Consequently, in 3 of Lemma 3.6, we record only the result that  $\mathcal{D}(\Pi_t)$  is equal to  $D_t$  in distribution (and not almost surely).*

**Lemma 3.8** *Let  $b_s$  be a finite bridge, and let  $V$  be a uniform random variable on  $[0, 1]$  which is independent of  $b_s$ . Let  $(H_i)$  be the (finite sequence of) holes of  $b_s$ , and let  $H = \cup_i H_i$ . Then there exists a uniform random variable  $U$  such that  $U < \mathcal{D}(b_s)$  if and only if  $V \notin H$ .*

PROOF: Note that if  $\mathcal{D}(b_s) = 0$  then the result is trivial, so assume  $\mathcal{D}(b_s) > 0$ . Let  $H_1, \dots, H_l$  be the holes of  $b_s$ , numbered such that  $\partial H_i < \partial H_j$  if  $i < j$ . By definition, the holes are disjoint. For each hole, let  $\delta H_j = \inf H_j$ , recall  $\partial H_j = \sup H_j$ , and set  $\partial H_0 = 0$ ,  $\delta H_{l+1} = 1$ .

For  $j = 0, 1, \dots, l$ , set  $I_j = [\partial H_j, \delta H_{j+1})$ . Define a function  $f : [0, 1] \rightarrow [0, 1]$  as follows.

$$f(y) = \begin{cases} \sum_{i=0}^{j-1} (\delta H_{i+1} - \partial H_i) + (y - \delta H_{j+1}) & \text{if } y \in I_j \\ \sum_{i=0}^l (\delta H_{i+1} - \partial H_i) + (y - \delta H_{j+1}) + \sum_{i=1}^{j-1} (\partial H_i - \delta H_i) + (y - \partial H_j) & \text{if } y \in H_j \\ 1 & \text{if } y = 1. \end{cases}$$

A graphical demonstration of  $f$  is given in Figure 5. It is easily checked that  $U = f(V)$  has the required properties.  $\blacksquare$

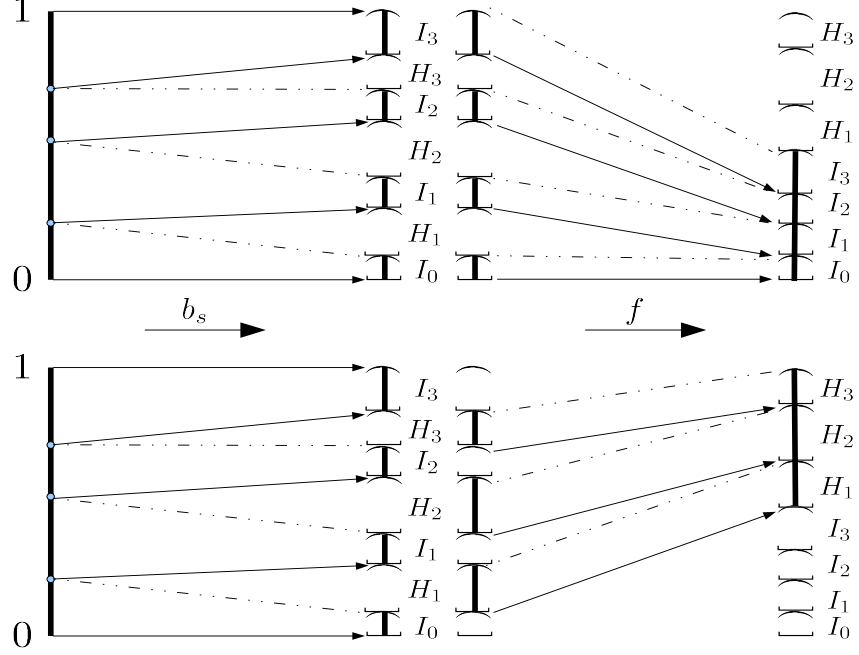


Figure 5: Diagram of  $f$ , from the proof of Lemma 3.8. The action of  $f$  is first to stack segments of dust on top of each other, and then stack the holes on top of that.

**Lemma 3.9** *Let  $m \in \mathbb{N}$ . Then there exists a sequence  $(U_i)_{i=1}^\infty$  of independent uniform random variables on  $[0, 1]$  such that the following holds. Almost surely, for all  $j = 1, \dots, m$ , the bridge  $B_{0,t}^{(N_m)}$  has at least*

$$\sum_{i=1}^{N_j} \mathbb{1}\{U_i \leq D_t\}$$

*holes of size at least  $\frac{D_t}{j}$ .*

PROOF: Note that a finite bridge has only finite many holes. Fix  $j \in \{1, \dots, n\}$ , and suppose  $H_1, \dots, H_l$  are the holes of some finite bridge  $b_s = b_{s_1, \dots, s_l}$ . Let  $x \in (0, 1)$ , and let  $V$  denote a uniform random variable (independent of  $b_s$  and  $x$ ) such that

$$b_x(a) = (1 - x)a + x\mathbb{1}\{V \leq a\}.$$

We consider what happens to the holes of  $b_s$  when we compose with  $b_x$  to form  $b_s \circ b_x$ .

The probability of  $V$  falling into the set  $\{\partial H_i; i = 1, \dots, l\}$  is zero, and henceforth we ignore this case (we claim only an almost sure result in the statement of this lemma). There are two further options:

(A) If  $V \in [0, 1] \setminus (\bigcup_i H_i \cup \{\partial H_i\})$ , then the holes of  $b_s \circ b_x$  are precisely the elements of the set

$$\{b_x(H_i); i = 1, \dots, l\} \cup \left\{ \left[ (1 - x)V, (1 - x)V + x \right) \right\}.$$

(B) If  $V \in H_k = [h_1^k, h_2^k]$ , for some  $k$ , then the holes of  $b_s \circ b_x$  are precisely the elements of the set

$$\{b_x(H_i); i = 1, \dots, k-1, k+1, \dots, l\} \cup \left\{ [b_x(h_1^k), b_x(h_2^k)] \right\}.$$

**Remark 3.10** *The reader might wish to refer back to Figure 4, where the two cases can both be seen. In that figure, (A) occurs when composing with  $b_{x_{\sigma(1)}}$  with  $b_{x_{\sigma(2)}}$ , and (B) occurs when composing  $b_{x_{\sigma(1)}} \circ b_{x_{\sigma(2)}}$  with  $b_{x_{\sigma(3)}}$ .*

In case (A), for all  $i = 1, \dots, l$  we say that the hole  $b_x(H_i)$  of  $b_s \circ b_x$  is a *child* of  $H_i$ . In case (B), for  $i = 1, \dots, k-1, k+1$ , we say also that the hole  $b_x(H_i)$  of  $b_s \circ b_x$  is a child of  $H_i$ . Additionally, in case (B) we say that  $[b_x(h_1^k), b_x(h_2^k)]$  is a child of  $H_k$ .

We label the following results for easy reference.

- (1) Every hole of  $b_s$  has a unique child amongst holes of  $b_s \circ b_x$ .
- (2) If a hole  $H'$  of  $b_s \circ b_x$  is a child of the hole  $H$  of  $b_s$ , then  $\mathcal{S}(H') \geq (1-x)\mathcal{S}(H)$ .
- (3) In case (A),  $b_s \circ b_x$  has one more hole than  $b_s$ , and in case (B)  $b_s \circ b_x$  has the same number of holes as  $b_s$ . In case (A), the new hole has size  $x$ .

The first of the above statements is essentially immediate, since  $b_x$  is a strictly increasing function with (by Lemma 3.4)  $\text{lip}^\downarrow(b_x) > 0$ . The third statement is immediate from (A) and (B). The second statement follows from Lemma 3.4, except in case (B) for the hole  $H = H_k = [h_1^k, h_2^k]$ . In this case  $H' = [b_x(h_1^k), b_x(h_2^k)]$ , and a direct calculation shows that

$$\mathcal{S}(H') = b_x(h_2^k) - b_x(h_1^k) = (1-x)(h_2^k - h_1^k) + x = (1-x)\mathcal{S}(H) + x.$$

By Lemma 3.2, the sum of the sizes of the holes of  $b_s$  is precisely  $1 - \mathcal{D}(b_s) = s_0$  where  $s_0 = 1 - \sum_1^\infty s_i$ . By Lemma 3.8, there is a random variable  $U \in \sigma(V)$  such that

$$\mathbb{1}\{U < \mathcal{D}(b_s)\} = \mathbb{1}\{V \text{ is not an element of a hole of } b_s\}. \quad (3.3)$$

Since  $V$  is independent of  $\mathcal{D}(b_s)$ ,  $U$  is independent of  $\mathcal{D}(b_s)$ . We record a fourth statement, which is also immediate.

- (4) Case (A) occurs precisely when (3.3) = 1, and case (B) occurs precisely when (3.3) = 0.

Now, fix  $m \in \mathbb{N}$ , and let  $\sigma = \sigma_{N_m}$ . Loosely speaking, we will apply the above reasoning iteratively along the composition  $b^{(N_m)} = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(N_m)}}$ , working from left to right. For each  $i = 2, \dots, N_m$ , at the  $i^{\text{th}}$  stage we set

$$b_s = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(i-1)}}$$

$$b_x = b_{x_{\sigma(i)}}.$$

We divide the remainder of the argument into two stages.

**CONSTRUCTION OF  $(U_i)$ :** Let  $(V_i)_1^\infty$  be a sequence of independent uniform random variables on  $[0, 1]$  which are independent of  $\sigma(\delta(M \cap (0, t) \times (0, 1)))$ , and use  $V_i$  as the uniform random variable (written as  $V$  above), at stage  $i$  of the iteration. It follows immediately that  $U_i$ , defined (at each step of the iteration) by (3.3), is also a sequence of independent uniform random variables on  $[0, 1]$ .

**CONSTRUCTION OF THE HOLES:** Fix  $j \in \{1, \dots, m\}$ . Recall that the sets  $\{x_i; i = 1, \dots, N_j\}$  and  $\{x_{\sigma(i)}; i = 1, \dots, N_j\}$  are in bijective correspondence. Hence, by definition of  $N_j$ ,



(5) There are at least  $N_j$  elements of  $\{x_{\sigma(i)}; i = 1, \dots, N_j\}$  for which  $x_{\sigma(i)} > 1/j$ .

Let  $i'$  be some  $i' \in \{1, \dots, N_j\}$  such that  $x_{\sigma(i')} > 1/j$ . If  $V_{i'}$  falls into a hole of

$$b_s = b_{x_{\sigma(1)}} \circ \dots \circ b_{x_{\sigma(i'-1)}}$$

(or equivalently, by (4), if  $U_i \leq \mathcal{D}(b_s)$ ), then, by (3), the new hole  $H$  of  $b_s \circ b_{x_{\sigma(i'')}}$  has size  $x_{\sigma(i')} \geq 1/j$ .

By (1), following the line of children of  $H$ , we reach a hole  $\tilde{H}$  of  $b_{0,t}^{(N_m)}$ , and by (2)

$$\mathcal{S}(\tilde{H}) \geq x_{\sigma(i')} \prod_{i=i'+1}^{N_m} (1 - x_{\sigma(i)}).$$

By (3.2),

$$\prod_{i=i'+1}^{N_m} (1 - x_{\sigma(i)}) \geq \prod_{i=1}^{N_m} (1 - x_{\sigma(i)}) = \mathcal{D}(B_{0,t}^{(N_m)}).$$

By Lemma 3.6,  $\mathcal{D}(B_{0,t}^{(N_m)}) \geq D_t$ , so in fact

$$\mathcal{S}(\tilde{H}) \geq x_{\sigma(i')} D_t \geq \frac{D_t}{j}. \quad (3.4)$$

To summarise, by (5), there are at least  $N_j$  possible distinct choices for  $i' \in \{1, \dots, N_m\}$  for which  $x_{\sigma(i')} > 1/j$ . By (4), when  $U_{i'} \leq \mathcal{D}(b_s)$ , each such  $i'$  gives rise to a hole of  $B_{0,t}^{(N_m)}$  which satisfies (3.4). By the uniqueness of children in (1), each  $i'$  for which  $U_{i'} \leq \mathcal{D}(b_s)$  gives rise to a unique hole of  $B_{0,t}^{(N_m)}$  (which, of course, satisfies (3.4)). Since the  $(U_i)$  are independent of  $N_j$ , this completes the proof.  $\blacksquare$

**Lemma 3.11** *For all  $\epsilon > 0$  and  $N \in \mathbb{N}$ , the set*

$$\mathcal{B}(N, \epsilon) = \{(s_i)_1^\infty \in \mathcal{P}_{\mathcal{M}}; \forall i = 1, \dots, N, s_i \geq \epsilon\}$$

*is a closed subset of  $\mathcal{P}_{\mathcal{M}}$ .*

PROOF: This follows from Lemma 2.4, since  $\mathcal{B}(N, \epsilon)$  is closed under pointwise convergence.  $\blacksquare$

PROOF: [Of Theorem 3.3.] The argument rests on an application of the Portmanteau Theorem, which can be found as Theorem 3.1 in Ethier and Kurtz (1986). The precise fact we require is the following.

- ( $\star$ ) Let  $(S, d_S)$  be a separable metric space, and let  $\mathbb{Q}_n, \mathbb{Q}$  (where  $n \in \mathbb{N}$ ) be probability measures on the Borel subsets of  $S$ . Let  $X_n$  and  $X$  be  $S$ -valued random variables with distributions  $\mathbb{Q}_n$  and  $\mathbb{Q}$ , respectively. Then,  $X_n \rightarrow X$  in distribution if and only if, for all closed subsets  $C$  of  $S$ ,  $\limsup_n \mathbb{Q}_n(C) \leq \mathbb{Q}(C)$ .

Of course, we will use the closed sets described by Lemma 3.11, and the convergence in distribution from Theorem 2.12.

Let  $\delta > 0$  and let  $N \in \mathbb{N}$ . By Lemma 3.6,  $D_t$  has the same distribution as  $\mathcal{D}(\Pi_t)$ , and by Theorem 1.7,  $\mathbb{P}[\mathcal{D}(\Pi_t) > 0] = 1$ . Hence, we may choose  $\epsilon > 0$  such that

$$\mathbb{P}[D_t \geq \epsilon] \geq 1 - \delta. \quad (3.5)$$

Since  $\nu([0, 1]) = \infty$  we have that  $\lim_{j \rightarrow \infty} N_j = \infty$ , almost surely. Hence, we may pick  $j \in \mathbb{N}$  such that

$$\mathbb{P}[N_j \geq N^2] \geq 1 - \delta. \quad (3.6)$$

We will use (3.5) and (3.6) frequently in the following estimates. For all  $m \geq j$ ,

$$\begin{aligned} & \mathbb{P} \left[ B_{0,t}^{(N_m)} \text{ has at least } N \text{ holes of size } \geq \frac{\epsilon}{j} \right] \\ & \geq \mathbb{P} \left[ B_{0,t}^{(N_m)} \text{ has at least } N \text{ holes of size } \geq \frac{D_t}{j} \right] - \delta \\ & \geq \mathbb{P} \left[ B_{0,t}^{(N_m)} \text{ has at least } \sum_1^{N_j} \mathbb{1}\{U_i \leq D_t\} \text{ holes of size } \geq \frac{D_t}{j} \right] \\ & \quad - \mathbb{P} \left[ \sum_1^{N_j} \mathbb{1}\{U_i \leq D_t\} \leq N \right] - \delta \\ & = 1 - \mathbb{P} \left[ \sum_1^{N_j} \mathbb{1}\{U_i \leq D_t\} \leq N \right] - \delta. \end{aligned}$$

To get from the penultimate to the final line of the above we use Lemma 3.9. Now,

$$\begin{aligned} \mathbb{P} \left[ \sum_1^{N_j} \mathbb{1}\{U_i \leq D_t\} \leq N \right] & \leq \mathbb{P} \left[ \sum_1^{N^2} \mathbb{1}\{U_i \leq D_t\} \leq N \right] + \delta \\ & \leq \mathbb{P} \left[ \sum_1^{N^2} \mathbb{1}\{U_i \leq \epsilon\} \leq N \right] + 2\delta \end{aligned}$$

Note that  $\sum_1^{N^2} \mathbb{1}\{U_i \leq \epsilon\}$  is just a binomial random variable with  $N^2$  trials and success probability  $\epsilon$ . Applying Hoeffding's inequality, we obtain

$$\mathbb{P} \left[ \sum_1^{N^2} \mathbb{1}\{U_i \leq \epsilon\} \leq N \right] \leq \frac{1}{2} \exp \left( -2 \frac{(N^2 \epsilon - N)^2}{N^2} \right) = \frac{1}{2} \exp \left( -2(N\epsilon - 1)^2 \right)$$

Collecting together, we have that for all  $m \geq j$ ,

$$\mathbb{P} \left[ B_{0,t}^{(N_m)} \text{ has at least } N \text{ holes of size } \geq \frac{\epsilon}{j} \right] \geq 1 - \frac{1}{2} \exp \left( -2(N\epsilon - 1)^2 \right) - 3\delta \quad (3.7)$$

Equation (3.7) is the crucial bound, and we can now move towards applying  $(\star)$ . Note that the right hand side of (3.7) does not depend on  $m$ .

Let  $\mathbb{Q}_m$  and  $\mathbb{Q}$  be the laws of  $\pi_t^{(N_m)}$  and  $\Pi_t$ , respectively. Then, by Lemma 3.2

$$\mathbb{P}\left[B_{0,t}^{(N_m)} \text{ has at least } N \text{ holes of size } \geq \frac{\epsilon}{j}\right] = \mathbb{Q}_m(\mathcal{B}(N, \epsilon/j)).$$

Similarly,

$$\mathbb{Q}(\mathcal{B}(N, \epsilon/j)) = \mathbb{P}\left[\Pi_t \text{ has at least } N \text{ blocks with asymptotic frequency } \geq \epsilon/j\right].$$

By Theorem 2.12,  $\pi_t^{(N_m)} \rightarrow \Pi_t$  in distribution, and by  $(\star)$ , and Lemma 3.11,

$$\limsup_{m \rightarrow \infty} \mathbb{Q}_m(\mathcal{B}(N, \epsilon/j)) \leq \mathbb{Q}(\mathcal{B}(N, \epsilon/j)).$$

Hence, by (3.7),

$$\begin{aligned} \mathbb{P}\left[\Pi_t \text{ has at least } N \text{ blocks with asymptotic frequency } \geq \epsilon/j\right] \\ \geq 1 - \frac{1}{2} \exp(-2(N\epsilon - 1)^2) - 3\delta. \end{aligned}$$

Recall that  $j$  depends on  $N$ , but  $\delta$  and  $\epsilon$  do not depend on  $N$ . Recall also that  $\epsilon$  depends on  $\delta$ , but  $j$  and  $N$  do not depend on  $\delta$ . From the above equation we have

$$\mathbb{P}\left[\Pi_t \text{ has at least } N \text{ blocks with asymptotic frequency } > 0\right] \geq 1 - \frac{1}{2} \exp(-2(N\epsilon - 1)^2) - 3\delta,$$

and letting  $N \rightarrow \infty$ ,

$$\mathbb{P}\left[\Pi_t \text{ has infinitely many blocks with asymptotic frequency } > 0\right] \geq 1 - 3\delta.$$

Since  $\delta > 0$  was arbitrary, we thus have

$$\mathbb{P}\left[\Pi_t \text{ has infinitely blocks with asymptotic frequency } > 0\right] = 1,$$

and the proof is complete. ■

### 3.3 An embedded $\Lambda$ -coalescent

In this section we show that, in a sense made precise by Theorem 3.12, the restriction of the  $\Lambda$ -coalescent to any infinite subset of its particles is also a  $\Lambda$ -coalescent. We then combine this with Theorem 3.3 to complete the proof of Theorem 1.8.

In this section we will consider  $\Lambda$ -coalescents with initial time  $t > 0$  (instead of  $t = 0$ , which was specified by Definition 1.1). We extend Definition 1.1 in the obvious manner.

Let  $\mathcal{F}_t = \sigma(\Pi_s; s \leq t)$ .

**Theorem 3.12** *Let  $T > 0$ , and let  $\mathcal{W} = \{w_i; i \in \mathbb{N}\}$  be a random,  $\mathcal{F}_T$  measurable, subset of  $\mathbb{N}$ , such that  $w_i < w_j$  if  $i < j$ . Suppose that  $\mathcal{W}$  contains precisely one element of each block  $b \in \Pi_T$ .*

*For any  $n \in \mathbb{N}$ , let  $\text{bl}_{\Pi_t}(n)$  be the (unique) block  $b \in \Pi_t$  such that  $n \in b$ . Define a process  $(\tilde{\Pi}_t)_{t \geq T}$  taking values in  $\mathcal{P}_{\mathbb{N}}$  by*

$$i \stackrel{\tilde{\Pi}_t}{\sim} j \iff \text{bl}_{\Pi_t}(w_i) = \text{bl}_{\Pi_t}(w_j). \quad (3.8)$$

*Then  $(\tilde{\Pi}_t)_{t \geq T}$  is a  $\Lambda$ -coalescent.*

In order to prove Theorem 3.12 we will need the following result.

**Theorem 3.13 (Pitman 1999)** *The  $\Lambda$ -coalescent is strongly  $\mathcal{F}_t$ -Markov.*

PROOF: [Of Theorem 3.12] Note that the map  $i \mapsto w_i$  is a bijection from  $\mathbb{N} \rightarrow \mathbb{N}$ , and hence  $\tilde{\Pi}_t$  is well defined. Define  $\eta : \mathbb{N} \rightarrow \mathbb{N}$  by  $\eta(w_i) = i$ . The initial state of  $\tilde{\Pi}_t$  is  $\tilde{\Pi}_T = \mathbb{1}_{\mathbb{N}}$ , since  $\mathcal{W}$  contains precisely one element of each block of  $\Pi_t$ .

Let  $\tilde{\Pi}_t^n = \iota_n(\tilde{\Pi}_t)$ . Suppose that  $\tilde{\Pi}_t^n$  has  $l \in \mathbb{N}$  blocks, and consider any subset  $\tilde{b}_1, \dots, \tilde{b}_k$  of distinct blocks of  $\tilde{\Pi}_t$ , where  $t \geq T$ . For each  $\tilde{b}_i$ , (by (3.8),

$$b_i = \bigcup_{n \in \tilde{b}_i} \text{bl}_{\Pi_t}(n)$$

is a block of  $\Pi_t$ , and for  $i \neq j$ ,  $b_i \cap b_j = \emptyset$ . Hence, by Theorem 3.13 (applied at time  $t$ ), and Definition 1.1, the rate at which the  $k$ -tuple of blocks  $\tilde{b}_1, \dots, \tilde{b}_k$  is

$$\lambda_{l,k} = \int_0^1 x^k (1-x)^{l-k} \nu(dx).$$

It is straightforward to see that this same rate of coagulation occurs independently for all  $k$ -tuples of  $\tilde{b}_1, \dots, \tilde{b}_l$ . By Definition 1.1,  $(\tilde{\Pi}_t)_{t \geq T}$  is a  $\nu$ -coalescent. ■

We are now in a position to prove our main result, Theorem 1.8.

PROOF: [Of Theorem 1.8] Let  $T > 0$  be deterministic and set

$$\mathcal{W} = \{\inf b; b \in \Pi_T, |b| > 1\}.$$

Note that, by Theorem 3.3,  $\mathcal{W}$  satisfies the conditions for Theorem 3.12. By Theorem 3.12, the corresponding coalescent  $(\tilde{\Pi}_t)_{t \geq T}$ , defined by (3.8), is a  $\nu$ -coalescent. By Theorem 1.7 (which applies since  $\mu^{-1} < \infty$  implies  $\mu^* = \infty$ ),  $(\tilde{\Pi}_t)_{t \geq T}$  does not come down from infinity. That is,

$$\mathbb{P} \left[ \forall t \geq T, |\tilde{\Pi}_t| = \infty \right] = 1.$$

By (3.8), each block of  $\tilde{\Pi}_t$  corresponds uniquely to an atomic block of  $\Pi_t$ , which implies that  $\mathbb{P}[\forall t \geq T, N_t^a = \infty] = 1$ . Since this holds for each  $T = 1/n$ , where  $n \in \mathbb{N}$ , we have that

$$\mathbb{P}[\forall t > 0, N_t^a = \infty] = 1.$$

This completes the proof. ■

## Acknowledgement

I am very grateful to both Alison Etheridge and Vlada Limic; the idea for this paper came out of conversation between the three of us. To the best of my knowledge, Theorem 1.8 was first conjectured by Vlada Limic, who also gave the first example/proof of a  $\Lambda$ -coalescent with  $N_t^a = N_t^s = \infty$  (in an unpublished note, using different methods to the proofs above).

## References

- J. Berestycki, N. Berestycki, and J. Schweinsberg. Small-time behavior of  $\beta$ -coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 44(2):214–238, 2008.
- N. Berestycki. *Recent Progress In Coalescent Theory*, volume 16. Ensaios Matematicos, 2009.
- J. Bertoin. *Random Fragmentation and Coagulation Processes*. Cambridge University Press, 2006.
- J. Bertoin and J. F. Le Gall. Stochastic flows associated to coalescent processes. *Probab. Th. Rel. Fields*, 126:261–288, 2003.
- M. Birkner, J. Blath, M. Capaldo, A. M. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger.  $\alpha$ -stable branching and  $\beta$ -coalescents. *Electron. J. Probab.*, 10(9):303–325, 2005.
- E. Bolthausen and A. Sznitman. On Ruelle’s probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197:247–276, 1998.
- P. Donnelly and T. G. Kurtz. Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab.*, 9(4):1091–1148, 1999.
- S. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.
- N. Freeman. Phase transitions in a spatial coalescent. <http://arxiv.org/abs/1109.4363>, pages 1–44, 2011.
- J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.
- J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.
- S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability*, 36:1116–1125, 1999.
- J. Schweinsberg. A necessary and sufficient condition for the  $\Lambda$ -coalescent to come down from infinity. *Electron. Comm. Probab.*, 5:1–11, 2000.
- J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106:107–139, 2003.